CLAIMS

What is claimed is:

1. An episode classification system including:

   a. a multitude of diagnosis records, each of said diagnosis records including:

      i. diagnoses information;

      ii. time of diagnoses information; and

      iii. patient information;

   b. a patient grouper for generating at least one patient group, each patient group generated by grouping patient records having similar patient information;

   c. a diagnosis grouper for generating at least one diagnosis group from a patient group, each diagnosis group generated by grouping patient records from a patient group that have similar diagnosis information;

   d. an episode analyzer including:

      i. a probability analyzer for performing probability calculations, each of said probability calculations capable of generating a probability value using at least two of said multitude of diagnosis records as input entries, said probability value representing the probability that said input entries belong to a single episode;

    ii. a episode grouper for grouping diagnosis records determined to belong to a single episode; and

    iii. a severity analyzer for performing episode severity calculations, each of said episode severity calculations capable of generating an episode severity value.

2. A episode classification system according to claim 1 wherein at least one of said diagnosis records is an anchor diagnosis record.

3. A episode classification system according to claim 1 wherein at least one of said diagnosis records is a trigger diagnosis record.

4. A episode classification system according to claim 1 wherein at least one of said diagnosis records is a stopping point diagnosis record.

5. A episode classification system according to claim 1 wherein said probability calculation:

    a. operates on a pair of diagnosis records; and

    b. is a function of:

i. a similarity value, said similarity value representing the similarity between said pair of diagnostic records; and

ii. a time between diagnosis value, said time between diagnosis value representing the time between said pair of diagnostic records.

6. A episode classification system according to claim 5 wherein said probability calculation includes a probability numerator divided by a probability denominator, said probability numerator set to said similarity value times a first constant, and said probability denominator set to the quantity of a second constant times said time between diagnosis value plus one.

7. A method for episode classification using a multitude of diagnosis records, each of said multitude of diagnosis records including: diagnosis information; time of diagnoses information; and patient information; including the steps of:

a. creating at least one diagnosis pair from said multitude of diagnosis records, each said diagnosis pair containing a unique combination of two diagnoses information;

b. for each said diagnosis pair, iteratively:

i. determining a co-occurrence value, said co-occurrence value being the number of unique patients for whom the two diagnoses contained in

each of said diagnosis pairs occurred within a co-occurrence window;
and

    ii. associating said co-occurrence value with each diagnosis information
contained in said diagnosis pair;

c. creating at least one patient group, each said patient group generated by
grouping said diagnosis records having similar said patient information; and

d. for each said patient group, iteratively:

    i. creating at least one diagnosis group, each said diagnosis group
generated by grouping said diagnosis records having similar said
diagnosis information;

    ii. for each said diagnosis group, iteratively adding a unique occurrence
identifier to said diagnosis information for each said diagnosis record;

    iii. creating at least one time between diagnosis pair from said diagnosis
records in said diagnosis group, each said time between diagnosis pair
containing a unique combination of two said diagnosis records;

    iv. for each said time between diagnosis pair, iteratively:

        1. setting a time between diagnosis pair value for each said
diagnosis pair equal to the absolute value of the difference
between said time of diagnoses information from each said
diagnosis record in said diagnosis group;

2. setting a score numerator equal to said co-occurrence value having the same combination of diagnosis information as said time between diagnosis pair value;

3. calculating a score for said diagnosis pair by dividing said score numerator by said time between diagnosis pair value; and

4. associating said score to said diagnosis pair;

v. setting a minimum score value equal to the minimum said score from the set of said scores associated to each of said diagnosis pairs in said patient group;

vi. setting a maximum score value equal to the maximum said score from the set of said scores associated to each of said diagnosis pairs in said patient group;

vii. setting a difference score value equal to difference of said maximum score value and said minimum score value;

viii. for each said diagnosis pair, iteratively:

1. setting a standardized score numerator value equal to said minimum score minus said score associated to said time between diagnosis pair;

2. setting a standardized score equal to said standardized score numerator divided by said difference score value; and

3. associating said standardized score to said diagnosis pair; and

ix. classifying each said diagnosis information into at least one episode using said standardized score.

8. A method according to claim 7 wherein said step of classifying each said diagnosis information into at least one episode includes the steps of:

a. flagging each of said diagnosis information in said patient group for analysis; and

b. until all diagnosis information in said patient group is analyzed, iteratively:

i. combining two of said diagnosis information in said patient group flagged for analysis which have the maximum said standardized scores not exceeding a preset cutoff into an episode record;

ii. creating a new diagnosis information, said new diagnosis information representing said diagnosis information in said episode record;

iii. calculating a new standardized score for said new diagnosis information by averaging the standardized score associated with each of said diagnosis information in said episode record; and

iv. de-flagging said diagnosis information in said episode record for further analysis.